StealthCrypter: A Secure File Encryption Tool with Authenticated Encryption and Secure Key Derivation

Syed Muhammad Shah
Department of Computer Science (IT, IOC)
Kohat University of Science and Technology (KUST)
Registration No: CS120222115
Kohat, Pakistan
linkedin.com/in/syed-muhammad-shah

***Abstract*—In the contemporary digital landscape, information security has transitioned from a technical requirement to a fundamental necessity [12]. This research-based case study explores the critical domain of file security, specifically focusing on the theoretical and practical application of symmetric key cryptography to ensure data confidentiality and integrity [9]. A primary motivation for this study stems from the observation that encryption is often implemented incorrectly, utilizing robust algorithms like AES in insecure modes of operation such as ECB or CBC without adequate integrity protections [8].**

To mitigate these vulnerabilities, this study advocates for the adoption of Authenticated Encryption with Associated Data (AEAD), specifically analyzing the Galois/Counter Mode (GCM) [1]. Furthermore, the research addresses the weaknesses of standard key derivation by evaluating Password-Based Key Derivation Function 2 (PBKDF2) with unique salts to neutralize brute-force attacks [2]. The findings synthesize complex cryptographic recommendations into a coherent, actionable framework for securing digital assets against an increasingly hostile cyber threat landscape.

***Index Terms*—AES-GCM, PBKDF2, Authenticated Encryption, Data Integrity, Information Security, PECA 2016, Cryptographic Engineering.**

I. Introduction

In the contemporary digital landscape, the security of information has transitioned from a technical requirement to a fundamental necessity [12]. The primary focus of this research is Cryptographic File Security, specifically the implementation of Authenticated Encryption (AE) to ensure both the confidentiality and integrity of data at rest [9].

A. Foundational Concepts

In the domain of information security, this study is defined through three foundational pillars:

- **Cryptography (Confidentiality):** The process of transforming plaintext into ciphertext to ensure that unauthorized individuals cannot interpret the content [8].
- **Data Integrity:** The assurance that data remains in its original state without unauthorized modification, insertion, or deletion [9].
- **Authenticated Encryption (AE):** A shared-key cryptographic form that simultaneously provides confidentiality and authenticity guarantees [1].

B. Legal Framework and National Initiatives

The Government of Pakistan has established a robust legal framework to combat electronic crimes [12]. Key initiatives include the Electronic Transactions Ordinance (ETO) 2002 and the Prevention of Electronic Crimes Act (PECA) 2016, which mandates the protection of sensitive digital assets [12]. Globally, standards like the General Data Protection Regulation (GDPR) and ISO/IEC 27001 have shifted the paradigm toward ”security by default”.

C. Key Research Contributions

The field of cryptographic engineering has been shaped by several key contributions:

- **McGrew & Viega:** Primary architects of the Galois/-Counter Mode (GCM), providing foundational mathematical proofs [4].
- **Moriarty:** Standardized the use of PBKDF2 for secure password-based key derivation [2].
- **Bellare:** Mathematically proved that ”Encrypt-then-MAC” logic is essential for secure AE schemes [9].
- **Gueron & Lindell:** Demonstrated that hardware acceleration (AES-NI) makes AES-GCM highly efficient [3].

D. Research Objectives

The primary objective of this case study is to bridge the gap between theoretical standards and practical file security. Specifically, this work:

1) Critiques legacy modes like AES-CBC and AES-ECB due to their lack of integrity protections [11].
2) Provides an architectural blueprint for ”StealthCrypter,” integrating PBKDF2 and AES-GCM [7].
3) Evaluates design resilience against modern threats using AI/ML analysis and mathematical modeling [5].

II. Literature Review

The literature surrounding file encryption and data integrity has evolved from focusing on pure confidentiality to a more holistic view including data authenticity and side-channel resistance.

A. Standardization and Theoretical Foundations

Foundational frameworks such as NIST SP 800-38D and RFC 8018 provide the compliance baseline for modern security tools [1]. NIST's recommendation of Galois/Counter Mode (GCM) is supported by McGrew and Viega, who provided the mathematical proofs for high-throughput security [4]. Furthermore, Bellare and Namprempre established the theoretical necessity of Authenticated Encryption (AE), proving that encryption without integrity is vulnerable to active attacks [9]. These studies confirm that a modern tool must utilize an AEAD scheme.

B. Performance and Implementation

Quantitative experiments by Gueron and Lindell demonstrated that AES-GCM is highly efficient on hardware-accelerated systems [3]. Sari et al. validated this approach for cloud storage using AES-256 for integrity [7]. Additionally, comparisons by Mahajan reinforce that AES remains superior to legacy algorithms like DES and RSA in terms of throughput and the avalanche effect [8].

C. AI and Emerging Threats

Recent studies explore the intersection of machine learning (ML) and cryptography. Barbosa et al. utilized ML to identify algorithms via ciphertext metadata, highlighting potential side-channel leaks [5]. Lee and Lee applied entropy analysis to distinguish legitimate encryption from ransomware, suggesting that modern tools must produce high-entropy, indistinguishable output to remain secure [6].

TABLE I
Summary of Key Literature and Contributions

No.	Study	Methodology	Contribution
1	Dworkin	Qualitative	Standardizes AES-GCM [1]
2	Moriarty	Qualitative	Defines PBKDF2 [2]
3	Gueron	Quantitative	Hardware optimization [3]
4	McGrew	Mathematical	GCM Security proofs [4]
5	Barbosa	AI/ML	Algo identification [5]
6	Lee	AI/ML	Entropy detection [6]
7	Sari	Mixed	Cloud framework [7]
8	Mahajan	Quantitative	AES vs DES vs RSA [8]
9	Bellare	Mathematical	AE foundations [9]
10	Krawczyk	Mathematical	HKDF paradigm [10]
11	IJFMR	Quantitative	ChaCha20 analysis [11]
12	Usman	Qualitative	Pakistan Law [12]
13	Bhale	Mixed	Hybrid storage [13]
14	IEEE	Mixed	Data sharing [14]
15	Rass	Game Theory	Defense strategies [15]

III. Research Methodology

To ensure a holistic analysis of the "StealthCrypter" project, this study employs a multi-faceted research methodology. Because information security is both a theoretical mathematical science and a practical engineering discipline, no single method is sufficient [15]. We utilized Quantitative, Qualitative, Mixed, Mathematical, and AI-based methods to analyze the problem from multiple perspectives.

A. Quantitative Method

We selected three quantitative studies to empirically measure the performance overhead of encryption.

- **Approach:** Benchmarks comparing AES-GCM, AES-CBC, and ChaCha20 were analyzed to determine the efficiency of the proposed tool on standard hardware [11].
- **Essence:** Focuses on numerical data such as throughput (MB/s), CPU cycles per byte, and memory usage.
- **Key Advantage:** Provides concrete evidence of efficiency, allowing for the mathematical prediction of encryption times for large datasets [3].

B. Qualitative and Legal Framework

This method focuses on the interpretation of text, policies, and standards.

- **Approach:** Analysis of NIST SP 800-38D, RFC 8018, and Pakistan's data protection laws (PECA 2016) to ensure legal compliance [12].
- **Key Advantage:** Ensures the project is built on a solid legal and ethical foundation, preventing technical designs that violate privacy regulations.

C. Mixed Methodology

This approach bridges the gap between abstract ideas and real-world application by validating theory through prototyping [13].

- **Approach:** Integration of separate components, such as AES encryption and SHA-256 hashing, into a single cohesive system.
- **Key Advantage:** Provides a realistic roadmap for development and highlights integration challenges, such as secure salt storage.

D. Mathematical Modeling and Simulation

Security is treated as a mathematical theorem that must be proven true through logic and probability theory [9].

- **Approach:** Analysis of probability bounds for brute-force attacks and mathematical guarantees of the GCM authentication tag [4].
- **Key Advantage:** Offers the highest level of assurance; if the mathematics is sound, the security is theoretically unbreakable within defined bounds.

IV. Results and Discussions

This section evaluates the efficacy of the StealthCrypter model by synthesizing qualitative standards, quantitative performance metrics, and mathematical security proofs.

A. Qualitative Analysis

The qualitative analysis confirms that modern security is driven by a combination of legal compliance and technical standardization [12]. The proposed design satisfies regulatory requirements such as PECA 2016 and GDPR, while adhering to NIST SP 800-38D standards for authenticated encryption. Using legacy modes like CBC without a separate Message Authentication Code (MAC) is now considered non-compliant with best practices.

TABLE II
QUALITATIVE COMPLIANCE ASSESSMENT

Parameter	Standard	Compliance Result
Legal Compliance	PECA 2016 (Pakistan)	Exceeded (via AES-256) [12]
Data Privacy	GDPR (Europe)	Satisfied
Mode of Operation	NIST SP 800-38D	Satisfied (AES-GCM) [1]
Key Strength	RFC 8018	Satisfied (PBKDF2) [2]

B. Quantitative Performance Metrics

Quantitative analysis focuses on the performance overhead of chosen cryptographic primitives. On hardware-accelerated processors, AES-GCM achieves significantly higher throughput than AES-CBC due to its ability to parallelize operations [3].

TABLE III
COMPARATIVE THROUGHPUT PERFORMANCE

Algorithm Mode	Throughput	Cycles/Byte	Latency (1GB)
AES-CBC	850 MB/s	1.2	1.18 s [3]
AES-GCM	2,800 MB/s	0.4	0.36 s [3]
ChaCha20-Poly1305	1,400 MB/s	0.8	0.72 s [11]

The "time-to-encrypt" for a 1GB file using AES-GCM is approximately 0.36 seconds, making the delay virtually imperceptible to the user.

C. Architectural and Integrity Validation

The "Onion Layer" approach (Password $\rightarrow$ Key $\rightarrow$ Authenticated Encryption) successfully mitigates common attack vectors [7]. Practical results prove that separating the "Authentication Tag" from the ciphertext allows the tool to detect file corruption immediately. If even a single bit is changed in a 1GB file, the decryption process will return a fail signal before any data is processed.

D. Mathematical Security Bounds

Mathematical analysis provides probabilistic security guarantees for the system:

- **Brute Force Probability:** With a 256-bit key, the keyspace size is 2^{256}. The probability of randomly guessing the key is $1/2^{256}$, which is mathematically negligible [8].
- **Collision Resistance:** The probability of successfully forging a 128-bit GCM tag is $1/2^{128}$, preserving mathematical integrity [4].
- **Key Derivation Cost:** Using PBKDF2 with 600,000 iterations reduces an attacker's guessing speed from 1 billion/sec to approximately 1,600/sec, effectively neutralizing brute-force threats [2].

V. CONCLUSION AND FUTURE DIRECTIONS

This study concludes that modern file security requires a "Security by Design" approach, where technical implementation is aligned with both legal standards and emerging cyber threats.

A. Limitations of the Study

While StealthCrypter presents a robust framework, several inherent limitations remain:

- **End-Point Vulnerability:** Encryption cannot protect data if the host machine is compromised by RAM-scraping malware or keyloggers [15].
- **Data Irrecoverability:** The absence of a "backdoor" means that if a user loses the salt file or password, the data is mathematically impossible to recover.
- **Metadata and Entropy:** Encrypted files exhibit high entropy, making them easily identifiable by AI-based detection tools [6].
- **Hardware Dependency:** Performance relies heavily on AES-NI instruction sets; legacy or low-power hardware may experience significant latency [3].

B. Future Research Directions

Future work for the StealthCrypter project should focus on expanding its functional and defensive capabilities:

- **Hybrid Systems:** Integration of Public Key Infrastructure (PKI) using RSA-4096 or ECC to facilitate secure file sharing [13].
- **Post-Quantum Cryptography:** Adoption of NIST-standardized algorithms like CRYSTALS-Kyber to ensure long-term resilience against quantum computing.
- **Steganography:** Combining AES encryption with steganographic carriers (e.g., JPEG or WAV files) to defeat AI-based entropy analysis.
- **HSM Integration:** Offloading key management to hardware security modules like TPM chips or YubiKeys to mitigate endpoint risks.

C. Conclusion

The StealthCrypter model demonstrates that confidentiality without integrity is insufficient for modern data protection. By transitioning from legacy CBC modes to Authenticated Encryption (AES-GCM) and mandating robust key derivation (PBKDF2), developers can build systems that satisfy both international standards (GDPR) and local legal requirements (PECA 2016) [12]. This research provides a blueprint for building trustworthy digital ecosystems in an increasingly hostile threat landscape.

REFERENCES

[1] M. Dworkin, "Recommendation for Block Cipher Modes of Operation: Galois/Counter Mode (GCM) and GMAC," NIST Special Publication 800-38D, 2007.

[2] K. Moriarty and B. Kaliski, "PKCS #5: Password-Based Cryptography Specification Version 2.1," RFC 8018, 2017.

[3] S. Gueron and Y. Lindell, "AES-GCM for Efficient Authenticated Encryption," in *Proc. 3rd Int. Conf. Real-World Crypto*, 2011.

[4] D. McGrew and J. Viega, "The Security and Performance of the Galois/Counter Mode (GCM)," in *Proc. INDOCRYPT*, 2004.

[5] F. M. Barbosa, A. R. S. F. Vidal, and F. L. de Mello, "Machine Learning for Cryptographic Algorithm Identification," *IEEE Latin America Transactions*, vol. 16, no. 7, 2018.

[6] S. Lee and K. Lee, "Machine Learning Based File Entropy Analysis for Ransomware Detection," *IEEE Access*, vol. 10, 2022.

[7] I. K. Sari et al., "Cryptographic Framework for Cloud-Based Document Storage Using AES-256 and SHA-256," *Journal of IT Kinesis*, 2025.

[8] P. Mahajan and A. Sachdeva, "A Study of Encryption Algorithms AES, DES and RSA for Security," *Global Journal of Computer Science and Technology*, 2013.

[9] M. Bellare and C. Namprempre, "Authenticated Encryption: Relations among Notions and Analysis," *Journal of Cryptology*, 2000.

[10] H. Krawczyk, "Cryptographic Extraction and Key Derivation: The HKDF Scheme," in *Proc. CRYPTO*, 2010.

[11] IJFMR, "A Python-Based Simulation and Security Analysis of ChaCha20 and AES," *Int. Journal for Multidisciplinary Research*, 2025.

[12] M. Usman, "Data Privacy Protection by Consumer Laws in Pakistan," *Islamabad Policy Research Institute (IPRI)*, 2024.

[13] P. Bhale, V. K. S. Gajendra, and M. Sujeet, "Secure File Storage Using Hybrid Cryptography," *Int. Journal of Innovative Science and Research Technology*, 2016.

[14] IEEE, "Secure Data Storage and Sharing Techniques for Data Protection in Cloud Environments," *IEEE Xplore*, 2024.

[15] S. Rass et al., "Game-Theoretic Cybersecurity: The Good, The Bad and The Ugly," *arXiv Preprint*, 2024.

"If you spend more on coffee than on IT security, you will be hacked. What's more, you deserve to be hacked."

— Richard Clarke

RESEARCH SUMMARY:

Enhancing Data Confidentiality and Integrity: A Case Study on Modern Cryptographic Standards and File Encryption Mechanisms